Rigid motion-resolved $B_1^+$ prediction using deep learning for real-time parallel-transmission pulse design

Alix Plumley$^1$ | Luke Watkins$^{1,2}$ | Matthias Treder$^3$ | Patrick Liebig$^4$ | Kevin Murphy$^2$ | Emre Kopanoglu$^1$

$^1$Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff University, Cardiff, United Kingdom
$^2$School of Physics & Astronomy, Cardiff University Brain Research Imaging Centre (CUBRIC), Cardiff, United Kingdom
$^3$School of Computer Science and Informatics, Cardiff University, Cardiff, United Kingdom
$^4$Siemens Healthcare, GmbH, Erlangen, Germany

**Purpose:** Tailored parallel-transmit (pTx) pulses produce uniform excitation profiles at 7 T, but are sensitive to head motion. A potential solution is real-time pulse redesign. A deep learning framework is proposed to estimate pTx $B_1^+$ distributions following within-slice motion, which can then be used for tailored pTx pulse redesign.

**Methods:** Using simulated data, conditional generative adversarial networks were trained to predict $B_1^+$ distributions in the head following a displacement. Predictions were made for two virtual body models that were not included in training. Predicted maps were compared with ground-truth (simulated, following motion) $B_1^+$ maps. Tailored pTx pulses were designed using $B_1^+$ maps at the original position (simulated, no motion) and evaluated using simulated $B_1^+$ maps at displaced position (ground-truth maps) to quantify motion-related excitation error. A second pulse was designed using predicted maps (also evaluated on ground-truth maps) to investigate improvement offered by the proposed method.

**Results:** Predicted $B_1^+$ maps corresponded well with ground-truth maps. Error in predicted maps was lower than motion-related error in 99% and 67% of magnitude and phase evaluations, respectively. Worst-case flip-angle normalized RMS error due to motion (76% of target flip angle) was reduced by 59% when pulses were redesigned using predicted maps.

**Conclusion:** We propose a framework for predicting $B_1^+$ maps online with deep neural networks. Predicted maps can then be used for real-time tailored pulse redesign, helping to overcome head motion–related error in pTx.

**Keywords**
deep learning, motion correction, parallel transmit, RF pulse design, ultrahigh field
1 INTRODUCTION

Parallel transmission (pTx) of RF pulses through independently controlled channels can help to overcome $B_1$ nonuniformity seen in the head at 7 T, particularly when tailored pulses are used. Tailored pulse design incorporates the measured transmit sensitivities ($B_1^+$) of each pTx channel, achieving a homogeneous flip angle across specified slices or regions. For optimal tailored pulse performance, the measured $B_1^+$ distributions must match those present at the time of pulse playout. However, channels’ electromagnetic fields (including $B_1^+$) and their interference patterns depend critically upon the object being imaged (i.e., the coil load), including its position, geometry, and composition.

Geometrical and compositional differences between human subjects are partly addressed in alternative, non-tailored approaches such as universal pulses (UPs), SmartPulse, and fast online-customized pTx pulses. Intersubject robustness is achieved by designing a UP (offline) to minimize error across a small database of representative subjects. An underlying assumption is that the range in head geometry and composition across human subjects is relatively constrained, implying that $B_1^+$ distributions are similarly constrained. The designed pulse (a minimum error solution for excitation over multiple $B_1^+$ distributions) is therefore assumed to work fairly well for any individual subject without the need for $B_1^+$ mapping. Plug-and-play usability of UPs in pTx has led to the method’s growing popularity.

However, the intersubject robustness of UPs comes at a cost to flip-angle uniformity. Tailored pulses typically yield lower normalized RMS error (nRMSE) of flip angle compared with UPs (7% vs 11% in Gras et al.). Additionally, the database approach is problematic in cases in which an individual is an outlier with respect to anatomies represented in the database. Moreover, these methods do not address the dependence of $B_1^+$ on load position, leading to unpredictable pulse performance in cases of different initial subject positioning and/or within-scan head motion. The former is often overlooked, whereas the latter is commonly reported. Large head movements (exceeding 20 mm/degree) often occur among certain clinical populations, elderly, and pediatric subjects. Because flip angle (and therefore the acquired signal) depends on $B_1^+$, displacements of approximately 5° have been found to cause an excitation error of 12%–19% (percent of target flip angle) when using pTx at 7 T, with larger movements causing larger flip angle–related artifacts.

A few approaches have been proposed to correct motion-related RF field changes. Faraj-Dana et al partially overcame motion-related effects on the (receive) $B_1$ field by simply reorienting coils’ measured sensitivity maps using a Euclidean transformation. Similarly, Wallace et al used radial basis functions to extrapolate channel sensitivities to voxel locations outside of the head, providing sensitivity information for all voxels in the FOV, regardless of head position. Extrapolated maps were used for retrospective correction. Neither approach considered dynamic motion-related field changes (e.g., changes in coil loading, shifting susceptibility gradients in tissue), as their effects were deemed minimal at 3 T. However, interactions between channels’ highly nonuniform transmit fields at 7 T, especially with pTx, indicate that dynamic motion-induced field changes cannot be overlooked. In contrast with these approaches, data-driven approaches inherently incorporate these changes.

Motion artifacts are often addressed through retrospective correction; however, this is problematic for several reasons. First, the issues described previously cannot be corrected retrospectively without motion-resolved $B_1^+$ maps, which are not available. Second, channels’ electric fields depend on the same factors, including load position. Specific absorption rate (SAR) distribution and associated tissue heating are therefore also sensitive to motion, and are especially so in pTx due to constructive interference between channels’ electric fields. Peak local SAR can exceed safety limits when head motion occurs in pTx simulations—a critical issue that cannot be addressed retrospectively. Conservatively bounded SAR estimates may be used, but this can prevent optimal imaging performance by limiting the RF power. In this study, the effect of motion on flip angle is the primary focus.

It is therefore desirable to overcome the motion dependence of tailored pTx pulse performance, and to do so using prospective techniques. Real-time pTx pulse design has been proposed as a solution, in which channels’ complex coefficients are continuously updated to counteract motion-induced sensitivity changes. Multispoke pTx pulses can be designed in less than 0.5 seconds, whereas 2D spatially selective spiral pulses can be estimated in about 9 ms using deep neural networks. With motion detection (e.g., Refs 22, 25, 34, and 35), channel updates could be determined by instantaneous head position, retaining flip-angle uniformity in cases of arbitrary and/or extreme motion. However, the required updates to channel coefficients depend on the motion-related field changes. Because real-time (i.e., motion-resolved) $B_1^+$ maps are not measurable, this requires that the relationship between head position and $B_1^+$ distribution to be characterized.

Deep convolutional neural networks have previously been used to estimate (non-pTx) $B_1^+$ distributions. In Wu et al., high-quality maps were predicted from reconstructed $T_1$-weighted images, removing the need for $B_1^+$ mapping, while still allowing retrospective correction of $B_1^+$ related artifacts in quantitative MRI. This approach
was limited to postprocessing; prediction quality deteriorated when $B_1^+$ was predicted directly from undersampled images. Abbasi-Rad et al used a convolutional neural network to reconstruct $B_1^+$ from a localizer scan for the purpose of SAR reduction through pulse scaling based on slice-wise $B_1^+$ magnitude; however, $B_1^+$ prediction quality was dependent on head position.37

In this work, we train a system of conditional generative adversarial networks38 to predict pTx $B_1^+$ distributions (referred to as $B_1$ maps) following simulated head motion, given the initial $B_1$ maps at the centered position as input. If used in conjunction with motion detection, this would constitute motion-resolved $B_1$-map estimation, and therefore permit real-time tailored pulse design. $B_1$-map prediction quality is assessed by comparison with ground-truth (simulation output) $B_1$ maps following motion. Furthermore, flip-angle distributions of multipole pTx pulses designed using network-predicted $B_1$ maps are compared with those produced by tailored pulses designed using the initial subject-specific $B_1$ maps alone. Finally, we also observe peak 10-g averaged local SAR for both pulses following motion.

2 | METHODS

2.1 | Simulations and data

Dizzy, Billie, Duke, and Ella (Figure 1A–D) of the Virtual Population39 (IT’IS, Zurich, Switzerland) were simulated with a generic 8-channel pTx coil in Sim4Life (ZMT, Zurich, Switzerland). Each model was simulated at one central, and 32 off-center, positions. Off-center positions included rightward 2, 4, 5, and 10 mm, posterior 2, 4, 5, and 10 mm, and all possible combinations thereof. These 29 positions are hereafter referred to as the R-P grid (Figure 1E). In addition, yaw 5°, 10°, and 15° positions were also simulated (Figure 1F). The Duke model was scaled to 90% of the original size, as the body and coil models intersected at some positions when the model was full-sized. To ensure consistent voxelization (and therefore consistent partial volume effects) in the body model across all simulated positions, the coil array was displaced rather than the body model. Simulations included the head, neck and shoulders,40 and were run at 295 MHz following coil tuning to this frequency. Simulation results were normalized to an accepted power of 1 W per channel beyond the input port to the coil elements, to override imperfections in coil matching and any positional dependencies. The simulations were manually checked for input impedance and reflection coefficient as well as field smoothness across positions.

Channels’ 3D $B_1$, electric field, current density, and SAR distributions were masked to exclude background (air) voxels and exported to MATLAB (The MathWorks, Natick, MA). To incorporate interactions between channels for local SAR evaluations, 10-g-averaged Q-matrices were calculated.28,41,42 Elements of the $8 \times 8$ Q-matrices were

$$Q_{ij}(r) = \frac{1}{2\rho(r)} \left[ J_{x,j}^H(r) E_{x,i}(r) + J_{y,j}^H(r) E_{y,i}(r) + J_{z,j}^H(r) E_{z,i}(r) \right],$$

where $\rho(r)$ is the tissue mass density (kg/m$^3$) in voxel $r$, $J$ is the complex current density (A/m$^2$); $E$ is the complex electric field (V/m); $x$, $y$, and $z$ are the three Cartesian axes; $i$ and $j$ are transmit channel indices; and $H$ denotes Hermitian transpose.

$B_1$ maps from 51 slices spanning a mid-axial slab with a thickness of 9 cm from the Duke and Ella body models (Figure 1C,D) were prepared for network training by interpolating to $256 \times 256$ in-plane resolution. The same preprocessing was applied to the Billie and Dizzy data, but at only six slice locations (Figure 1A,B). Magnitude and phase data were separated and normalized between 0 and 1, where 1 corresponds to the maximum magnitude across all channels, slices, and body models, and to $2\pi$ for phase. Random offsets were applied to phase maps so that the phase wrap boundary did not occur at the same location across slices. $B_1$ maps were input to networks as individual axial slices with size $256 \times 256 \times 8$, where the third dimension is channels. Corresponding $B_1$-map slices before (input) and after (ground truth) a given displacement formed the networks’ input-target pairs. Note that inputs are not necessarily at the centered position (explained later in Section 2.2).

2.2 | Neural networks and network training

Models were implemented in TensorFlow 2.343 using Python 3.7. Network architecture is summarized in Figure 2. Except where specified, network hyperparameters were the same as those used in the Pix2Pix conditional generative adversarial network.44 The generators were U-Net45 models with eight convolutional (encoding) and eight deconvolutional (decoding) layers linking the input and output (predicted) $B_1$ maps, each followed by rectified linear unit activation layers. Filters were $4 \times 4$ for magnitude and $8 \times 8$ for phase. Although comprehensive hyperparameter optimization was beyond the scope of this project, during initial testing it was found that phase networks benefited from the large receptive field of $8 \times 8$ filters. Conversely, magnitude networks generated
smoother maps when more filters were used. To avoid increasing the number of trainable parameters, filters were smaller for magnitude. The number of filters (initially 128 for magnitude, and 64 for phase) increased to a maximum of 1024 (512 for phase) for the middle layers, and stride size was 2. Filters were split into eight groups to facilitate simultaneous processing of all pTx channels. For phase, batch normalization was applied at all layers except the first convolution layer. For magnitude, removing batch normalization resulted in a smoother training curve and higher-quality estimated maps. Skip connections joined each convolution layer to the symmetric de-convolution layer for network stability. The network was regularized through dropout layers following each of the first three deconvolution layers (rate = 0.5).

In contrast to encoder–decoder models that typically rely on minimizing L1 loss between predicted and target images, generative adversarial networks include an additional loss term, which helps to reduce blurring often seen with L1 loss alone.44 This is provided through a second convolutional neural network—the discriminator—which is trained to distinguish between generator-predicted and ground-truth distributions. The input $B_1$ maps, concatenated with either ground-truth or generator-predicted $B_1$ maps, served as input to the discriminators, which consisted of five convolution layers. The discriminators used...
leaky rectified linear unit activation layers ($\alpha = 0.3$) as recommended in Radford et al.\textsuperscript{46} Filter size was the same as that for the generators, and convolution stride was 2 except for the final two layers, where it was 1. A single 2D distribution of probability (entropy) values was output.

The overall conditional generative adversarial network loss function can be expressed as

$$\text{Loss} = \arg\min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G). \quad (2)$$

where $G$ denotes the generator; $D$ is the discriminator; and $\lambda$ (set to 100) is a scaling parameter acting on the L1-norm between generator-predicted and ground-truth maps. The first term can be further described as

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{B_{1gt}, B_{1predicted}} \left[ \log D \left( B_{1gt}, B_{1predicted} \right) \right] + \mathbb{E}_{B_{1gt}, B_{1initial}} \left[ \log \left( 1 - D \left( B_{1gt}, G \left( B_{1gt}, B_{1initial} \right) \right) \right) \right], \quad (3)$$

where $B_{1gt}$ are the ground-truth displaced $B_1$ maps; $B_{1predicted}$ are the generator-predicted displaced $B_1$-maps; and $B_{1initial}$ are the pre-displacement $B_1$ maps (network input).

The effect of head motion on $B_1$ depends on the displacement type (i.e., direction, magnitude).\textsuperscript{12,13} Because data-driven approaches assume that all input-target pairs share a common underlying mapping, separate networks were trained for different displacement types (e.g., rightward vs posterior). Head motion was discretized into large (5 mm) and small (2 mm) displacements in rightward (R) and posterior (P) directions to cover the R-P grid. Additional networks were trained for $5^\circ$ yaw rotation. Separate networks were trained for magnitude and phase, yielding a total of 10 networks.

The Duke and Ella data were used for training. All available relative displacements were included. For example, to train the R5 mm network, such as
Networks were tested using the Billie and Dizzy data at six slice locations (Figure 1A). For Billie, different slices were used compared with those used for early stopping (Dizzy was not involved in the training process at all). Like training, testing was conducted for all available examples of each displacement, yielding test data sets of between 6 and 72 slices. In addition to the positions listed in section 2.1, Billie and Dizzy models were simulated at three combined yaw-rightward positions to test performance for motion involving both rotation and translation. Because networks were only trained for five displacements but evaluated at 35 positions, networks were cascaded where necessary. Starting with the center position’s $B_1$ maps as input, generators were run sequentially, with the output of one generator used as input to the next, until the desired evaluation position was reached. For example, R5 mm, R5 mm, and P2 mm networks were cascaded for evaluation at the (R10, P2 mm) position. Finally, the Billie model was also simulated at inferior 5, 10, and 15 mm to investigate error for through-plane motion.

Predicted $B_1$ maps were exported to MATLAB. Voxels with $<1\%$ of the maximum magnitude were smoothed with a Gaussian kernel. Corresponding magnitude and phase network outputs were subsequently combined to form complex predicted maps ($B_1\text{predicted}$).

The $B_1\text{predicted}$ quality was assessed through voxel-wise correlation (using MATLAB’s corrcoef function) and prediction error between predicted and ground-truth maps at each position. These values were compared with error and correlation following head motion (i.e., between the two simulated maps). Calculations were performed on the tissue-masked region, with the outermost two voxels excluded to avoid artificial amplification of error due to partially filled voxels. Prediction error for magnitude and phase distributions were assessed through nRMSE and L1 norm, respectively, as follows:

$$nRMSE_{\text{prediction}} = 100 \times \sqrt{\frac{1}{N_r} \sum_{r=1}^{N_r} \frac{\|B_1gt(r) - B_1\text{predicted}(r)\|^2}{\|B_1gt(r)\|^2}}$$  

$$L1_{\text{prediction}} = \frac{1}{N_v} \sum_{r=1}^{N_r} \|e^{iB_1gt(r) - B_1\text{predicted}(r)}\|$$

where $j$ is $\sqrt{-1}$; and $N_v$ is the number of voxels in a slice, indexed by $r$. Motion-induced error was calculated analogously, but substituting $B_1\text{initial}$ for $B_1\text{predicted}$ in Equations 4 and 5.

### Pulse design and analysis

Outputs from the R-P grid positions were further processed to assess whether predicted maps were of sufficient quality to be used for tailored pTx pulse design. Five-spoke excitation pulses were designed using a small tip-angle spatial domain method, and two approaches were compared in terms of their performance following motion within the R-P grid. A schematic of the process is shown in Figure 3. First, a conventional tailored pulse (pulse$_{\text{initial}}$) was designed using the subject-specific $B_1$ maps at the initial position ($B_1\text{initial}$). A uniform magnitude target excitation profile (target flip angle = $70^\circ$) was specified for pulse$_{\text{initial}}$. Pulse coefficients were optimized iteratively to minimize magnitude error, whereas the profile’s phase was relaxed. The resultant complex profile was used as the target profile for a second pulse (pulse$_{\text{re-designed}}$), which was designed without phase relaxation (because magnitude and phase distributions need to be consistent across positions to ensure data consistency for motion occurring mid-acquisition). Pulse$_{\text{re-designed}}$ was designed using the network-output $B_1\text{predicted}$ (the proposed approach).

Both pulses (pulse$_{\text{initial}}$ and pulse$_{\text{re-designed}}$) were subsequently evaluated using the ground-truth $B_1$ maps at the displaced position ($B_1\text{gt}$) to quantify motion-induced effects on the conventional approach, and improvement provided by the proposed method. Their flip-angle distributions were compared with that of pulse$_{\text{initial}}$ without motion in terms of nRMSE, expressed as percent target flip angle as follows:

$$nRMSE_\theta = 100 \times \sqrt{\frac{1}{N_r} \sum_{r=1}^{N_r} \frac{\|\theta_{\text{displaced}}(r) - \theta_{\text{initial}}(r)\|^2}{\theta_t}}$$
where $\theta_{\text{initial}}$ is flip angle without motion; $\theta_{\text{displaced}}$ is that following motion (with either $\text{pulse}_{\text{initial}}$ or $\text{pulse}_{\text{re-designed}}$); and $\theta_t$ is the target flip angle. The nRMSE for $\text{pulse}_{\text{initial}}$ without motion (i.e., the “gold standard”) was also calculated by substituting $\theta_{\text{displaced}}$ for $\theta_t$ in Equation 6.

Peak local SAR (psSAR) of both pulses was also evaluated using the 10-g averaged Q-matrices at each position. Because psSAR sensitivity to motion has been reported to exhibit no slice dependence, SAR was evaluated at four target imaging slices (out of the six used for pulse design) (Figure 1A).

### RESULTS

#### 3.1 B$_1$ prediction quality

B$_1$ maps were predicted by networks in about 14 ms using the same PC as used for training. Example B$_1$ maps, motion-induced error, and prediction error are shown for a 5 mm displacement in Figure 4. Motion-induced error (averaged across channels) for this example was 15.1% (magnitude) and 4.9° (phase), whereas mean prediction error was 3.2% (magnitude) and 3.5° (phase).

Figure 5 shows a summary of error and correlation coefficient for magnitude and phase at each evaluated displacement (averaged across Dizzy and Billie models, slices, channels, and initial positions). Position dependence of prediction quality was minimal compared with motion-related error, as seen by the reduced gradient with respect to displacement norm in all cases. Dizzy and Billie models yielded very similar prediction quality (Supporting Information Figure S1).

Mean motion-induced magnitude error scaled linearly with displacement magnitude at about 3% per millimeter (or 3.2% per degree of rotation), compared with 0.36% per millimeter (0.27% per degree) for prediction error. Prediction error was lower than motion-related error in 99.8% of translation, and 90% of rotation evaluations. Figure 6A shows B$_1$ magnitude nRMSE for magnitude for all slices and channels for 10 example displacements. Motion caused a worst-case magnitude error of 117% following a displacement of R20, P10 mm, whereas maximum prediction error was 33% (at the y15°, R4 mm position). Worst-case prediction error from the R-P grid was lower (20% at the R20, R10 mm position).

Example magnitude correlations are shown in Figure 7A. The lowest observed correlation coefficient between $\text{B}_{1_{\text{initial}}}$ and $\text{B}_{1_{\text{gt}}}$ magnitudes was 0.79 following a y15°, R4 mm displacement. Correlation between $\text{B}_{1_{\text{predicted}}}$ and $\text{B}_{1_{\text{gt}}}$ did not fall below 0.96.

Motion-induced error and correlation were observed to be slice-dependent and channel-dependent (i.e., the error depended on the displacement relative to each channel’s location, as expected). The $\text{B}_{1_{\text{predicted}}}$ quality was similar across channels, as demonstrated by the strong correlation...
across all channels in Figure 7A. However, prediction error was somewhat slice-dependent, with the most inferior slice locations yielding the highest prediction errors (slice information not shown).

Phase error for 10 example displacements are shown in Figure 6B. For phase, maximum observed prediction error (57°) was similar to maximum motion-related error (55°). These worst cases arose in the Dizzy model; for the Billie model, maximum prediction error (29.4°) was lower than that caused by motion (44.2°). Furthermore, prediction error was lower than motion-induced error for phase in 68% and 66% of translation and rotation evaluations, respectively (including both models).

Yaw rotation caused substantially higher error than axial translations; for the R-P grid, maximum prediction error (29.4°) was lower than that caused by motion (44.2°). Furthermore, prediction error was lower than motion-induced error for phase in 68% and 66% of translation and rotation evaluations, respectively (including both models).

Mean phase correlation coefficient between predicted and ground-truth maps was higher than (or very similar to) that between initial and ground-truth maps for all displacements. Phase correlation examples are shown in Figure 7B. Correlation coefficient between $B_1^{\text{predicted}}$ and $B_1^{\text{initial}}$ exceeded that between $B_1^{\text{predicted}}$ and $B_1^{\text{gt}}$ in 69% of cases.

### 3.2 Parallel-transmit pulse performance

Subsequent analyses were conducted using the Billie model with the R-P grid data only. Five-spokes pTx pulses designed using $B_1^{\text{initial}}$ ($\text{pulse}_{\text{initial}}$) yielded uniform flip-angle profiles (mean nRMSE ~1%) without motion. However as expected, uniformity was lost following axial translation. Pulses were about 7.7 ms long.

Figure 8 shows that flip-angle nRMSE for pulse$_{\text{initial}}$ was strongly position-dependent, reaching a maximum of 76% following a displacement of R20, P5 mm. Conversely, pulses redesigned using $B_1^{\text{predicted}}$ ($\text{pulse}_{\text{redesigned}}$) produced much improved flip-angle profiles when evaluated at the displaced position, yielding nRMSE of 14% for the same displacement. Maximum pulse$_{\text{redesigned}}$ nRMSE was 17% (at the R2, P10 mm position), whereas this error value was exceeded by
FIGURE 5 Error (nRMSE for magnitude, L1 norm for phase) and correlation coefficient ($\rho$) shown for magnitude and phase, averaged over Dizzy and Billie body models, channels, slices, and initial positions for each evaluated displacement. Translational displacements (the R-P grid) are shown in the large panels, while rotations (yaw) and combined rotation-translations (yaw plus a 4-mm translation) are shown in the smaller panels below (for the purpose of the x-axis, the amount of yaw rotation is treated as magnitude displacement; for example, yaw $5^\circ$ plus 4-mm translation is shown at x = 6). The effects of motion are shown in purple, while network-related prediction error is shown in yellow. The SD is shown as shaded regions for magnitude but is omitted for phase for clarity, as values were similar.
Pulse \(\text{initial}\) (i.e., without any correction) after displacements of just \(\geq 5\) mm. The largest errors occurred in inferior slice locations for both pulses (slice information not shown). Maximum motion-related error in the excitation profile’s phase (110.4°) was reduced by 7.8° when using pulse \(\text{redesigned}\).

Figure 8B shows flip-angle profiles for both pulses following several example displacements. Supporting Information Figure S2A also shows flip-angle nRMSE for nine example displacements. It should be noted that flip-angle uniformity for pulse \(\text{redesigned}\) could be further improved by including phase relaxation in the design (as was done for pulse \(\text{initial}\)); however, this would permit excitation phase to vary throughout the scan, causing reconstruction inconsistencies.

3.3 Cascading

The B\text{1\ predicted} quality remained high when networks were cascaded multiple times; however, there was a weak linear relationship between prediction error and motion magnitude. To investigate the impact of cascading on prediction quality, we ran secondary analyses for displacements of R0, P10 mm, R-2, P10 mm, and R-5, P10 mm using only the P2 mm network for the posterior component. Running the 2 mm network five times (i.e., four cascades) led to approximate average increases in magnitude and phase error of 1.2% and 1.2°, respectively, compared with running the 5 mm network twice (one cascade). There was also reduced flip-angle uniformity compared with using the 5 mm network. Nevertheless, Figure 9 shows that
motion-induced error was appreciably reduced using either approach.

3.4 Specific absorption rate

In addition to flip angle, SAR was also evaluated for the R-P grid positions. Following motion, psSAR produced by pulse\textsubscript{redesigned} was lower than that of pulse\textsubscript{initial} in 89% of cases. For pulse\textsubscript{initial}, motion caused psSAR to increase (relative to that without motion; psSAR\textsubscript{center}) in 72% of evaluations. When pulse\textsubscript{redesigned} was used instead, psSAR increased relative to the centered case in only 16% of the cases. The psSAR for nine example displacements is shown in Supporting Information Figure S2B.

Figure 10A shows worst-case psSAR for each evaluated position (relative to psSAR\textsubscript{center}). Figure 10B shows the same information, separated by slice location. Motion-related SAR change was similar across slices, whereas benefits offered by pulse\textsubscript{redesigned} were most consistently seen in slices 1 and 6 (furthest from mid-axial locations). Pulse\textsubscript{redesigned} yielded lower psSAR than pulse\textsubscript{initial} following large displacements, but sometimes resulted in higher psSAR for small displacements, especially in slice 4 (mid-axial). In the worst observed case, pulse\textsubscript{redesigned} yielded a 3.1-fold increase in psSAR (slice 4 at the posterior 5-mm position), whereas pulse\textsubscript{initial} increased by a maximum of 3.3 fold (following the largest axial displacement). It should be noted that SAR was not used as a constraint in either pulse’s design.

4 DISCUSSION

As far as we know, this study is the first to demonstrate motion-resolved B\textsubscript{1} map prediction in pTx. We successfully
PLUMLEY et al. trained a system of deep neural networks to estimate $B_1^+$ sensitivity distributions following simulated in-plane head motion. Predicted $B_1^+$ maps were of sufficiently high quality to be used for tailored pTx pulse design, and resulting pulses outperformed conventionally designed tailored pulses under conditions of head motion.

Across the four R-P grid magnitude networks, prediction error was $4.5\% \pm 1.5\%$ (mean $\pm$ SD) of the ground-truth $B_1$ magnitude (tested individually without cascading and error calculated according to the definition in Abbasi-Rad et al$^{37}$). This is in line with expectations based on previous work, in which magnitude $B_1$ maps were predicted with mean error of $9.5 \% \pm 7.1\%$. The experimental and 3D nature of data in Abbasi-Rad et al$^{37}$ may have caused the slightly higher error. We also observe similar $B_1^{\text{predicted}}$ correlation coefficients to those reported in Wu et al$^{36}$ (~0.99), which was retained even when our networks were cascaded multiple times. Yaw rotation caused higher error than translational motion, and prediction error was generally higher than that of translations. We attribute this to the much smaller training database used to train yaw networks.

It was previously reported that excitation profile nRMSE increases by approximately 2.4% per millimeter of axial displacement in simulations,$^{12}$ and our results (2.7% per millimeter) are in good agreement. A 12%–22% flip-angle error was observed in vivo following motion in the range of about 5–16 mm in Kopanoglu et al.$^{12}$ and we observed an error of about 11%–35% for a similar range of motion. Error was reduced to approximately 8%–10% using the proposed approach in our study. For larger movements, the benefit of pulse redesign using predicted maps was even greater.

There was some slice dependence for both $B_1^{\text{predicted}}$ quality and flip-angle error, with the highest errors observed for inferior slice locations. This is in line with previous research, in which higher motion sensitivity was observed for inferior slices passing through the temporal lobes and the cerebellum.$^{12}$ Inferior slices yield lower field magnitude, and $B_1^{\text{predicted}}$ quality was lower in regions with very low field magnitude, which may explain the residual slice dependence in predicted maps. An alternative explanation is that there were fewer inferior slices in the training data set compared with mid-axial slices, which could have resulted in better training for mid-axial slices.

Considering that the R5 mm training data included positions up to just 10 mm along the rightward axis, it is noteworthy that the R5 mm network was able to extrapolate beyond this by successfully cascading four times to estimate the fields at the R20 mm position. We attribute this to the global normalization, conducted over all data sets. Results for large displacements could likely be further improved by including extreme positions in the training data set.

Magnitude networks consistently outperformed those of phase. Prediction error exceeded motion-induced error for phase in about a third of evaluations. Improvements to excitation phase were evident but modest. It has been

**FIGURE 8** Excitation profile results for five-spoke pTx pulses following head motion. (A) Mean flip-angle nRMSE (above) and phase RMSE (below) for excitation profiles, averaged over slices and initial positions for each evaluated displacement. Excitation pulses were five-spoke pTx pulses designed using either the initial position (pulse$_{\text{initial}}$) or predicted (pulse$_{\text{redesigned}}$) $B_1$ maps. The SD is shown as shaded regions for magnitude but is omitted for phase for clarity, as values were similar. (B) Example flip-angle profiles produced by pulse$_{\text{initial}}$ at the initial position, by pulse$_{\text{initial}}$ at the displaced position, and by pulse$_{\text{redesigned}}$ at the displaced position.
previously acknowledged that phase changes due to motion are difficult to model, and other work on motion-related field changes, including $B_1$ prediction, often neglect phase altogether. Most of the phase-prediction error occurred at phase-wrapping boundaries. This was somewhat reduced by applying random phase offsets to pairs of input $B_1$ maps, but it was not eliminated. In terms of pulse performance, these small, local errors are likely to have less impact than the global changes caused by motion, which could potentially accumulate when channels are superposed. Incidentally, when we instead trained networks using unwrapped phase data, the error in $B_1_{\text{predicted}}$ was globally higher and yielded inferior results compared with wrapped data. In contrast, the local prediction errors seen with wrapped data are not structured. We believe that the increase in error seen for unwrapped data was due to the increased dynamic range of unwrapped data, meaning that relative changes due to motion were smaller following normalization.

One limitation of this study is that it deals with simulated data only because of the lack of models for the RF coils at the institution due to proprietary information. In DeepQSM, training data were solely synthetic; the ground truth consisted of overlapping cubes and spheres with known (simulated) susceptibility values. This was convolved with a forward dipole kernel to create the corresponding input. Networks were able to resolve high-quality susceptibility maps for human brains, despite only being...
trained on simple geometric shapes. Similarly, Meladio et al demonstrated successful in vivo validation following training with synthetic data. We believe that the method proposed here would be similarly generalizable if a realistic RF coil model (i.e., a model of the coil to be used) is used for simulations. Moreover, using simulated training data avoids the requirement for choreographed in vivo head movement to be replicated precisely across several subjects to create the training data sets, which would be practically infeasible. Aside from the initial (measured) input $B_1$ map, networks automatically output maps in patient coordinates, making online registration unnecessary. However, motion tracking is needed to determine which network(s) are required, and for online corrections to gradient waveforms to update the imaging volume as in Zaitsev et al.
The minimum motion resolution we consider here is 2 mm. Smaller movements could remain problematic for quantitative MR protocols that rely on signal changes on the scale of 1%–4% of the total signal. Finer discretization is possible through simulations of smaller displacements. The most appropriate motion discretization will depend on the user’s primary aims. While using finer discretization and cascading more (i.e., $5 \times 2$ mm vs $2 \times 5$ mm) did result in slightly higher error here, motion-induced error was still largely ameliorated.

In this study we considered only positive displacements (e.g., rightward, but not leftward). Because the trained networks are direction-specific, the same networks cannot handle negative displacements. This is due to the nature of inverting a deep neural network (not possible, as reversing convolutional layers results in a highly underdetermined problem). However, the training of networks for positive displacements does not limit generalizability, and additional networks can similarly be trained for negative displacements. In fact, by training a network for positive and negative displacements for each of the 6 degrees of freedom of motion, all rigid-body motion could be covered.

We focused on in-plane motion (axial translations and yaw rotation), as motion-induced error was observed to be more spatially varying. Although some motion-induced $B_1$ error was observed in peripheral slices following inferior translation (see Supporting Information Figure S3), the error was relatively lower, and importantly, more global (i.e., spatially smoothly varying) within the slice. We attribute this to the fact that relative tissue-channel distances remain constant for through-slice translation. This means that simpler correction methods (e.g., slice-dependent pulse scaling) could feasibly be used to counteract the $B_1$ effects of through-plane motion.

In contrast to $B_1$, through-plane motion was shown to be more disruptive to $B_0$ than within-plane motion. Although $B_0$ off-resonance can be incorporated in tailored pulse design, this cannot currently be updated in real time with pTx, as it increases the degrees of freedom to be optimized in the pulse, pushing redesign times beyond practically feasible TR values. Instead, motion-related effects on $B_0$ can be corrected retrospectively, such as using data-driven coefficients to link motion with field changes. Alternatively, real-time $B_0$ shimming may be possible with multcoil shim arrays by predicting $B_0$ field changes due to motion in a manner similar to the method proposed here.

The SAR observations reported here are incidental. The focal point of this study was to develop a method to accurately estimate $B_1$ maps following motion. Using SAR as a pulse design constraint would trade flip-angle homogeneity for reduced SAR, thereby overshadowing $B_1$ quality. Hence, SAR was not used as a design constraint. Although motion-related SAR increase was generally lower for pulses redesigned with predicted $B_1$ maps, it was higher for a minority of cases. Motion sensitivity of SAR in pTx has previously been reported to be similar across axial slices, and we also did not observe clear slice dependence for SAR motion sensitivity. However, we did observe that inferior-most and superior-most slices benefited from the proposed approach more consistently than mid-axial slices. Nevertheless, the overall improvement offered by pulse redesigned, especially for larger displacements, is promising for future development of this approach. Neural networks have previously been used to predict $B_1$ maps for the purpose of SAR reduction. This was achieved through slice-wise pulse scaling based on a predicted 3D $B_1$ magnitude. The entire 9-cm axial slab could be predicted within approximately 0.8 seconds, permitting concatenation into pseudo-3D $B_1$ maps that could feasibly be used for pulse scaling or similar SAR management here. Pulse scaling based on $B_1^+$ cannot guarantee SAR compliance, as $B_1^+$ does not necessarily reflect electric-field distributions. However, SAR compliance could be ensured if (3D) electric fields were also predicted.

5 | CONCLUSIONS

We have demonstrated a framework for a deep-learning approach for motion-resolved $B_1^+$ estimation in pTx. Estimated maps can be used for real-time tailored pulse redesign, yielding homogeneous flip-angle profiles in cases of head motion. Importantly, networks can be run sequentially to predict $B_1$ maps following arbitrary displacements comprising multiple directions. Here, error was reduced for 35 displacements using networks trained for just five displacements. Our findings represent one potential avenue toward user-friendly, optimized pTx at 7 T.

CONFLICT OF INTEREST

Patrick Liebig is an employee of Siemens Healthcare.

ORCID

Alix Plumley & https://orcid.org/0000-0002-5404-7491
Luke Watkins & https://orcid.org/0000-0001-9286-3134
Patrick Liebig & https://orcid.org/0000-0001-7342-3715
Emre Kopanoglu & https://orcid.org/0000-0001-8982-4441

TWITTER

Alix Plumley & @Alix_992
REFERENCES


**SUPPORTING INFORMATION**
Additional supporting information may be found in the online version of the article at the publisher’s website.

**FIGURE S1** (A) B1 prediction quality (error and correlation) for the Billie model only. (B) B1 prediction quality (error and correlation) for the Dizzy model only. (Figure 5 displays the same evaluations averaged over the two body models.) Data from the Billie model were used for validation during the training process, whereas the Dizzy model was not involved at all during training.

**FIGURE S2** (A) Raincloud plots showing flip-angle normalized RMS error (nRMSE) (percent of target flip angle) for nine example displacements using pulse_initial (orange) and pulse_redesigned (dark blue). Asterisks indicate the number of network cascades required for evaluation. The light blue–shaded region shows maximum prediction nRMSE across all 35 evaluated displacements (consistent across panels). (B) All specific absorption rate (SAR) evaluations for the same nine example displacements. Vertical axes show relative peak local SAR (psSAR), calculated as psSAR as a factor of that without motion using pulse_initial (psSAR_center). The green-shaded region shows psSAR at or below psSAR_center (i.e., indicating that psSAR did not increase following motion). Neither pulse included SAR constraint in the design.

**FIGURE S3** Voxelwise correlations between initial (B1_initial) and ground-truth (B1gt) B1 maps, to compare motion-related error following through-plane (A) and within-plane (B) displacements. Results are shown for two positions and three slice locations (1 = least superior, 6 = most superior). Note that the large inferior displacement (15 mm) is greater than the large rightward displacement (10 mm), and yet error remains lower due to the through-plane nature of the movement.